

Noname manuscript No.
(will be inserted by the editor)

A New Emotional Robot Assistant that Facilitates Human Interaction and Persuasion

Jaime A. Rincon · Angelo Costa · Paulo Novais · Vicente Julian · Carlos Carrascosa

Received: 02 Feb 2018 / Revised: 02 May 2018 / Accepted: 26 May 2018

Abstract The development of robots that are truly sociable requires understanding how human interactions can be applied to the interaction between humans and robots. A sociable robot must be able to interact with people taking into account aspects like verbal and non-verbal communications (emotions, postures, gestures). This work presents a social robot which main goal is to provide assistance to older people in carrying out their daily activities (through suggestions or reminders). In addition, the robot presents non-verbal communications like perceiving emotions and displaying human identifiable emotions in order to express empathy. A prototype of the robot is being tested in a daycare centre in the northern area of Portugal.

Keywords Social Robots; Emotional Models; Ambient Assisted Living

1 Introduction

Globally, the elderly population is increasing, according to demographic projections (United Nations, Department of Economic and Social Affairs, Population Division, 2015). According to the OMS, the amount of people aged over 60 is expected to double between 2000 and 2050 (United Nations, Department of Economic and Social Affairs, Population Division, 2017). As less developed countries start to evolve, this trend is onset immediately (United Nations, Department of Economic and Social Affairs, Population Division, 2017). A common societal issue that emerges from a rapid elderly population growth is the exponential demand or care (medical and otherwise).

The common human resources needed are healthcare professionals, including formal and informal caregivers. Which currently are in short supply, and there is no indication of increase of the caregivers in the foreseeable future. The lack of these professionals represents a serious issue. One way of overcoming this issue is the usage of technology that empowers people to overcome problems that they encounter on their daily lives. Various areas can be of help to the elderly, like Social Robotics (Broekens

Universitat Politècnica de València. D. Sistemas Informáticos y Computación jrincon, vinglada, carrasco@dsic.upv.es · Centro ALGORITMI, Escola de Engenharia, Universidade do Minho, Braga acosta, pjon@di.uminho.pt

et al, 2009), Virtual Assistants (Krämer, 2008), Artificial Intelligence (Lupiani et al, 2017), etc.

One of the most used technology is humanoid and animal-like robots, which show that people socialize easier with them due to their aspect. These robots can help on-site caregivers by playing a part in nursing the elderly or be used as a social too by providing advices and maintaining a conversation, keeping an active company for the elderly (Broekens et al, 2009; Feil-Seifer and Mataric, 2005). These robots are expected to communicate in a natural and intuitive way. Furthermore, the price of social robots is decreasing and one can be bought or build for under 500 €, which means that even people with small economic power are able to have one at their home.

In terms of adoption, people are increasingly become familiarized to have some type of robot or autonomous entities at their home, e.g., Roomba vacuum cleaner and Amazon's Alexa. Although these robots are limited in their reach and features, they are slowly setting the path so more complex and advanced systems can be used in a home environment without being rejected by the residents.

Currently, there are already some robots present in key areas that give the people an insight of the future, thus gaining trust from the people they interact with. These robots can be found working in hospitals, elderly residencies, etc. They are able to initiate an interaction with humans, showing proactive features, thus being more similar to human-like interaction.

Although we can observe advanced robot systems like the Pepper robot (Soft-Bank, 2018) and PARO robot (Wada et al, 2003) displaying human-like traits, they lack necessary tools that are required for challenging environments such as nursing homes or assisting physically/cognitively challenged people. It is necessary to introduce new features that empower the robots on these type of environments. For instance, one of the features can be non-verbal communication, e.g., display of emotions, human gestures, and postures. Additionally, the usage of anthropomorphic features increase the acceptance of robots by humans, that is why social robots typically present rounded faces, eye, mouth and ears. We believe that these features are key for improving the acceptance of robots, and build trust and cooperation.

Emotion capture and display is important to show underlying psychological characteristics (making them more human-like). Humans are able to empathize and create emotional bonds, thus it is important that the robots tap into that process to communicate better. Furthermore, perceiving the emotional state of the people can help the robot make decisions that are adjusted to each person personality.

There are models of emotion and personality that are quantified, being mathematically representable, thus they are able to be used in computer systems. A few examples are OCC (Ortony, 1990), PAD (Mehrabian, 1997), Circumplex Model (?) and OCEAN (McCrae and John, 1992). However, these models only represent the emotion of a single individual and not of a group of individuals. Since the goal of social robots is to be able to cater to several people, they are on their own inadequate to be singularly used. Each individual is different and responds differently to the environment stimuli, which modifies their emotional state. Therefore, it is necessary to introduce new models that are able to represent a group social emotion (Rincon et al, 2017), with the goal of being incorporated in robots.

Apart from this topic, new developments in robotics area should try to endow the robots with decision support procedures and persuasion procedures that in combination provide more and better information to the users, and, in this particular

case, to elderly people. The persuasion tries to improve user engagement to the system for the selection of activities that potentially best suits to the users' interests. In this sense, robot might effectively use behaviors such as vocal and bodily cues, to persuade users (Chidambaram et al, 2012). Other studies have shown how the robot gender also influences on human behavior (i.e. men were more likely to donate money to the female robot) (Siegel et al, 2009). On the other hand, the social feedback provided by a robot can also be employed in order to obtain persuasive effects over users (Ham and Midden, 2014).

In this paper we present our social robot *EmIR* that has as goal to provide assistance to the elderly performing their daily activities, presenting human-like features like perceiving emotions of a group of people and displaying human identifiable emotions. Moreover, our social robot makes use of an e-Health platform (CAL) which integrates a persuasive module that makes use of argumentative techniques that map the reasoning procedures that physicians and caregivers follow to recommend activities to patients.

This paper is structured as follows. Section 2 analyzes previous works in the area of assistants robots. Section 3 presents in detail the main functionalities of the robot with particular emphasis on the emotion detection and the user interaction. Section 4 describes the hardware employed to build the robot. Section 5 illustrates the functionalities of the robot through an example. Finally, some conclusions are presented in Section 6.

2 Related Work

Recently it is observable the increase on interest in assistive robots and their development (Martinez-Martin and del Pobil, 2017b,a). Current sociological shift demands technological solutions that are able to interact with elderly people. Thus, the sudden increase of the number of robots available that have as goal solving this sociological issue. Furthermore, they are increasingly refined and advanced in terms of aspect and features, being more human-like in both fields.

Next we present three projects that represent these advancements and that are in the same domain as our project.

2.1 Pepper

The Pepper robot (SoftBank, 2018) is a personal assistant robot that aims to keep company, interact with human beings and perform very limited tasks. This robot has a humanoid aspect (with head, torso and arms) while having wheels as the propeller system for moving around easily. The success of this robot is astounding and was well-received by the population, being adopted at homes and in commerce (for publicity purposes).

The aim is to be a companion to people (mostly directed to elderly people) by interacting with them, reading their emotions to create an empathic connection, and sustaining an intelligible conversation (retrieving facts and information from the internet).

In terms of features it is composed by a computer system and an array of sensors. Using cameras and microphones it interprets the environment, being able to detect

shapes and speaking patterns. Furthermore, with these sensors it is able to move around a space without colliding into other objects or people.

The sympathetic aspect is the success to this specific robot, as it is limited in its goal. Its sensor array is imprecise and is only able to detect 4 emotions (joy, sadness, anger and surprise) on exceptional environmental conditions (lightning and sound). These conditions are not found in normal environments, thus Pepper is unable to detect accurately an emotion. To overcome this issue, it relies on artificial intelligence to assert a probable emotion. In terms of tasks, due to the fragile hardware and small frame, it is only able to perform gestures and carry lightweight objects, like glasses of water. Lacking a complete hand structure it is unable to grasp objects.

2.2 PARO

Another type of robot is PARO (Personal Assistant RObot) (Wada et al, 2003), which has the shape of an animal, in this case a baby of harp seal. This robot was built to be a companion to children and elderly, prompting a positive emotional response due to its visual aspect.

The robot has a soft fur covering it and inside tactile, vision, audition and posture sensors. With these sensors it is able to perceive direct interaction, like touching and speaking. It is very limited in terms of processing, but it compensates these limitations by moving its flaps and tail when a person interacts with it. It understands basic language commands and is able to emit some sounds when activated.

Although PARO is limited in terms of functionality, people that have interacted with it report increased levels of happiness, better mood and improvement in terms of communication (Wada and Shibata, 2008). Furthermore, caregivers have reported that the use of PARO has a positive impact on the environment, being adamant on its use on everyday activities and even more for therapy (Niemela et al, 2016).

2.3 AIDO

The AIDO is a robot from InGen Dynamics (Dynamics, 2018) with a semi-human shape and with the aim of being an home assistant. While the focus is not directly assisting the elderly people, its features and functionality can be used to assist people on their daily tasks. Its main function is to interact with people and provide information when requested. It has no means of physically interact with people (it does not have any type of arms or similar tools) but is able to move around the house (currently resorting to wheels).

Their current prototype uses screens and projectors to show the robots emotions based on the person input and interaction. It is unable to detect emotions and the display of its own is based on vocal perception. Apart from that, it is able to process vocal commands and is able to provide facts and information (via a connection to the internet).

From these project we are able to observe that while the physical aspect, and its underlying impact (as displayed by PARO), are very advanced and the target of a careful development, they lack in terms of functionality. The actions that makes one human and their emotions are not explored by these projects and we believe that

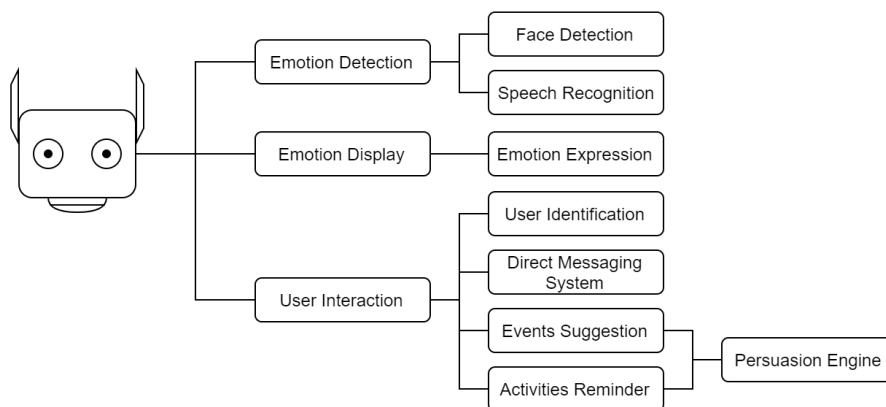


Fig. 1 Compact services and frameworks representation used by the platform.

after the initial impact of the visual aspect is what makes them really human-like. We, as humans, strive for human empathy, and is what our project aims for.

3 Emotional Robot Assistant Description

This section describes the proposed social robot, called **EmIR (Emotional Intelligent Robot)**, which has been developed to provide assistance to the elderly performing their daily activities. A prototype of this robot is being validated in a real scenario, in a daycare institution in the north of Portugal.

The main features of the robot are the following: perceives the emotions of each individual and also calculates the social emotion of a group of people; displays human identifiable emotions according to the emotional states of the users; assists users recommending healthy activities and also, remembering when to do them; and finally, persuades users by giving them arguments to justify recommended actions.

The functionalities of the proposed robot are divided into three main modules which provides the services that our robot uses to perceive and interact with the environment (see Figure 1). The first two modules are the Emotion Detection and the Emotion Display modules which are in charge of detecting, processing and displaying emotions. These two modules are explained in greater detail in sections 3.1 and 3.1.2. On the other hand, the other module is the User Interaction module (explained in section 3.3) which has been implemented using the CLA (Cognitive Life Assistant) framework. The CLA framework improves the user experience and gives a fluid visual interface to the robot.

3.1 Emotion Detection

The interaction is an important part whitening human society, through this, is possible express the necessities, give orders, etc. This information is communicated using different channels and in each one of them different messages can be found. Humans in the course of his life, can learn to understand these messages presented into these signals (i.e. emotions, voice level, body gesture, etc.). This learning process associated

to this information can help to understand the behaviour, to suggest something that helps to modify it. These processes, that are easy to learn and to understand for humans, are not so easy for a social robot, since the social robot needs different elements (hardware and software) that help to sense, understand and reproduce some elements of human behaviour. The task to reproduce human behaviour through a social robot incorporates different elements, seeing that, the social robot needs to reproduce the identification of a person, of an emotion, speech recognition, body gesture, etc. To do this, it is necessary to supply different elements such as sensors that allow to the social robot to acquire information. Nevertheless, as sensors are not enough, it is necessary to use some Artificial Intelligence (AI) tools that help to analyze and learn from the information acquired by them. Within these tools, we can find algorithms to face detection, face recognition, natural language processing, emotion detection, etc. In this section, we present the different AI tools used into the social robot that help it to recognize a person, emotions, speech interaction and activity suggestion.

3.1.1 People Identification and Emotion Classification

A social robot aims to interact with people, as well as, to know who the person with whom it interacts is. In the literature, we can find different techniques to identify people such as neural networks (Caramazza et al, 2017) and deep learning (Castro et al, 2017). In these techniques it is necessary to have a large set of data with different facial positions to train the machine learning model. The process to train this model takes a long time and in some cases is annoying. This process should be quick and comfortable, trying to reduce the learning time and not disturb the user.

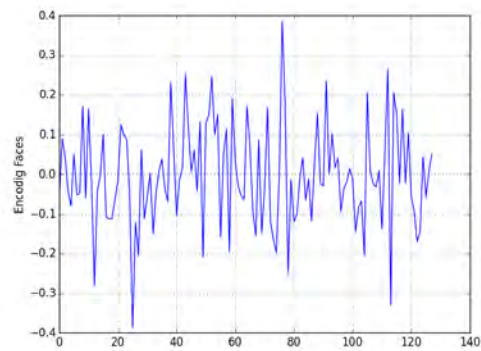
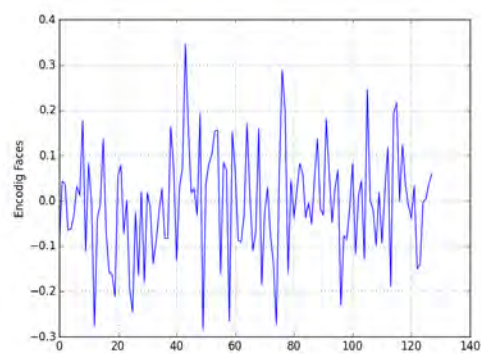
Our social robot reduce the training time using a technique of mapping the human faces. Faces are converted into vectors of 128 elements, where each picture of the same person is mapped. When the robot takes an image, it extracts the face and converts it into a vector with a dimension of 128 descriptors. The social robot uses this technique to identify the people around him/her. Figure 2 shows a graphic of the coding of the image. In this Figure it is possible to see a certain difference between users, these vectors represent the footprint of the face to recognize.

To perform the person identification, the system calculates the distance between these two vectors. If the difference is within an error threshold, the system returns the name of the user it has recognized (Figures 3).

The use of this method allows to reduce the training process, since only with an image the robot could identify a person. However, this technique is used only to identify persons, in the case of emotions this process does not yield good results. This happens because the emotional expression may vary in different individuals. For this reason, it was decided to use convolutional neural networks to classify human emotions.

Convolutional Neural Networks (Krizhevsky et al, 2012) are very similar to ordinary Neural Networks. They are made up of neurons that have learnable weights and biases. The whole network still expresses a single differentiable score function: from the raw image pixels on one end to class scores at the other.

The robot is able to classify 7 emotions: fear, angry, upset, happy, neutral, sad and surprised. To train the model, the KDEP dataset (Calvo and Lundqvist, 2008) , which consists of a total of 980 images, was used. In this dataset the 70 actors, 35 women and 35 men, represent the 7 different emotions. Each image has a resolution

(a) **A** user coding(b) **B** user coding**Fig. 2** Encoded user vectors.

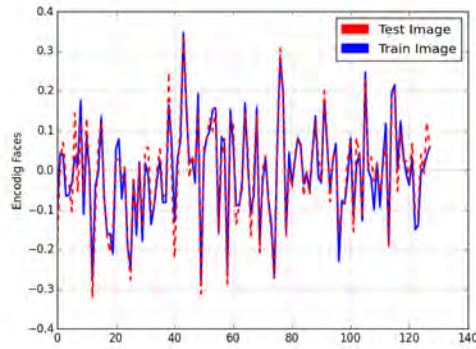
of 562 pixels wide and 762 pixels high. Before performing the training, each image is pre-processed to detect and extract the face. Once extracted, a color transformation is performed in grayscale. Finally, it is necessary to resize the image to 48x48, and to train the model using Keras ¹ and Tensor Flow ².

Figure 4 shows the Model Accuracy in the Train Data and Validation Sets, seeing how the test graph follow the training graph. This behaviour allows the pressure level between the training data set and the test data to be determined. Figure 5 shows the Model Loss in the Training and Validation Data Sets, this graph shows the validation process between the training model and our test data.

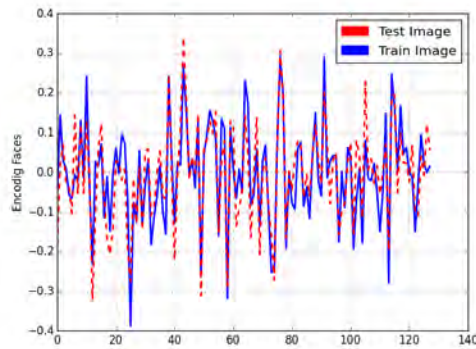
The model training is expensive and uses a large part of the machine's resources and the robot does not have this capability. So, training and testing is done on a server. This server incorporates training and testing algorithms, and offers these services that can be used by the robot or other entities to identify a person or classify an emotion. This web service offers the services of people identification,

¹ <https://keras.io/>

² <https://www.tensorflow.org/>



(a) Comparison coding training image with test image User **A**.



(b) Comparison coding training image with test image User **B**.

Fig. 3 Comparison coding training image with the image taken.

emotion rating and social emotion calculation (if the robot is in front of more than two people) (Rincon et al, 2016a). The following subsection explains in detail the main functionalities of the web service.

3.1.2 Emotional Web Service for Social Robots

In order not to load processes inside the robot and distribute the information, we decide to build a Web Service. This Web Service implements the techniques of person identification and emotions' recognition explained in the previous section. So, it offers the robot the ability to identify the person, to recognize their emotional state and to calculate the Social Emotion if it is in front of more than two people. These elements help the robot to make a better decision making, when it comes to interacting with the human. When it knows who the person is and their emotional state, a series of services that helps the person are offered.

However, if the robot does not recognize the person standing in front of it, the robot initiates an interaction process in order to know who it is. To do this, the

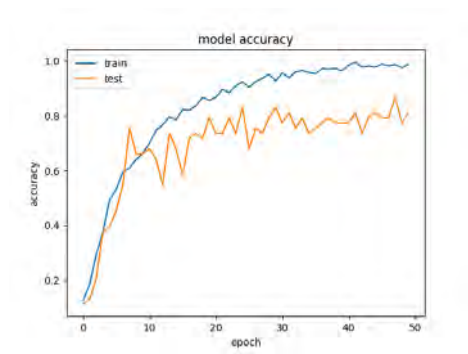


Fig. 4 Model Accuracy on Train and Validation Datasets

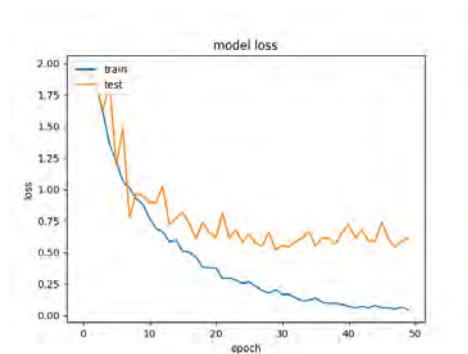


Fig. 5 Model Loss on Training and Validation Datasets

robot has a series of dialogues, these dialogs facilitate interaction with the human. These dialogues allow to create a communication with the human, to make a set of questions. These questions are what humans normally use to interact with a person they do not know.

Once the robot captures the name of the person, this name is sent to the Web Service so that the name is associated with the image and, therefore, the robot learns this new face. This service can be used by several robots or other entities that are within the environment. It is important to note that the robot learns the faces separately, that is, if there is more than one face that the robot does not know, it has to use this service as many times as faces exist.

Emotion detection is performed by the Web Service request to analyze the image sent. If there is only one person inside the image, the service returns the classified emotion and the robot responds to the emotion it has classified. But if the robot detects more than one face, the service returns the emotion of each one of the individuals present within the image, as well as the Social Emotion of the group. This Social Emotion allows us to know what is the emotional state of the group in order to act if this Social Emotion is too far away from the goal one. How the social emotion of a group of people is calculated, is out of the scope of this paper and can be consulted in (Rincon et al, 2016a).

3.2 Emotion Display

As we have mentioned, emotions play an important role for assistant robots, as people behaviour changes according to their emotions. So it is very important to detect people emotions, but in the same sense is very important to improve the interaction with humans, that the robot will be able to express its own emotional state so it would be easily integrated in an interaction with humans. For this reason, the robot presented in this paper is able to express its own emotional state. In order to express emotions, the robot uses its sensors to perceive its surroundings, and use this information to generate its emotional state. The robot express its emotional state in two different ways:

- By a color: the robot has two different LED emitters in its cheeks. These LEDs will light with a different color depending on the robot emotional state. This color is calculated according to an easy conversion between PAD and RGB.
- By expressing the emotion by means of its face: the robot will move its eyebrows and mouth to express one out of seven different emotions (according to the discretization of its emotional state). This movements will be done to reach one of the seven different robot face drafts shown in Figure 6.

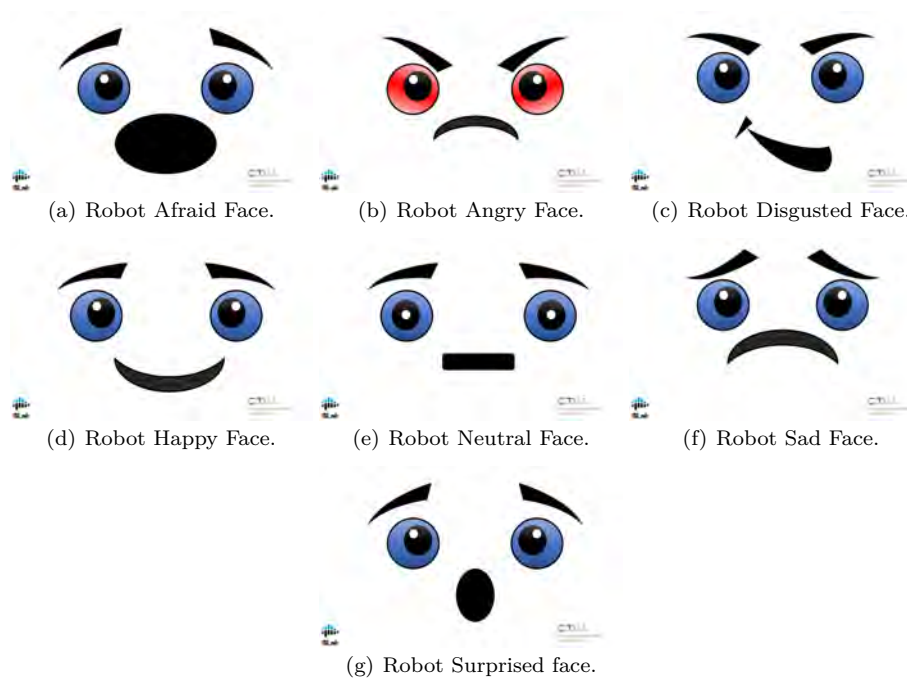


Fig. 6 Robot Face Draft.

Using these two methods, the robot lets the humans know that it also feels sad, happy, surprised or annoyed. It helps to generate a friendly bond between the human

being and the robot, so humans may help the robot to feel good, just as the robot does when the human is not. It is a way to create an acceptance feeling by humans.

3.3 User Interaction

The *User Interaction* is supported by the CLA (Cognitive Life Assistant), being its features seen in Figure 1 connected to the *User Interaction*. The aim is to provide the robot with features that can be used to improve the user experience and give a fluid visual interface.

CLA is a cognitive assistant (Costa et al, 2017b), and is designed to be a tool for the elderly, aimed at helping them perform their daily tasks and promoting active and independent aging. At its core, CLA revitalizes one's memory by prompting reminders about activities, information about how to perform them and suggesting activities that have a positive effect on the health condition (both physical and cognitive).

CLA is composed by several modules (Costa et al, 2012), of which 5 are specifically used on this robot (Figure 1), being: user identification, events suggestion (Rincon et al, 2016b), activities reminder (Rincon et al, 2016b), persuasion engine (Costa et al, 2016) and the direct messaging system. They give the ability of interacting with the users and provide useful information and advice at the right time.

Each module/engine main feature is:

- **User Identification** This module keeps the information of each user about all personal and medical aspects. The CLA has the goal of improving the health condition, thus the ability of knowing what activities are appropriate for each user is required. Furthermore, the persuasion module needs information about the users' age, gender and family, friends and physicians contacts, so it is able to correlate the suggestions with the personal environment.
- **Events Suggestion** This module selects and suggests events to the user based on its profile and medical condition. The objective is to engage the users in active aging, by suggesting events that keep them physically and cognitively active, e.g., reading a book, a walk in the park, a bridge game with friends. Additionally, if friends of the user are registered in the CLA group events can be suggested. The suggestion is accompanied by a date which all users are available. This module is composed by the following elements:
 - A database with tables of the users, activities, relatives and friends. This database keeps all persistent information of the relation between the users' likes and activities that they can perform, the relation between the users and a complete list of activities.
 - The scheduler algorithm. This algorithm creates a composed list of activities and prioritizes them according to the users' preferences and needs. Then, when it is required by the CLA or in a timed manner, it prompts the users an activity from the composed list (which follows a selection process that considers the number of times that an activity was suggested, the number of acceptances by the user, and considerations like the level of intensity or if it is recommended by an expert) and saves it to the users' calendar if accepted. Augmenting this algorithm is the possibility of the suggested activity being shared by other users.

- The secondary scheduler algorithm. This algorithm carries the previous algorithm and verifies if the activity has the “shared” flag as true, it searches for friends that may be invited and invites them, if this activity can be scheduled on the users’ invited free time.

In the case of the EmIR, when a user is detected, this module operates in a timed fashion (configurable to each user) and when the module is triggered an event is suggested and presented to the user for his/her acceptance, activating the persuasion engine to provide persuasion arguments that legitimate this suggestion and tries to nudge the user into acceptance. The persuasion engine is explained in greater detail below.

- **Activities Reminder** This module presents information about the activities present in the users’ calendar. The goal is to inform about the next activity and medication intake. The robot is able to take the information provided by this module and relay it to the users (vocalizing it or making head gestures) for critical tasks (health-related ones). Furthermore, the CLA is able to save and present additional information for each activity, e.g., what public transportation can be used to reach a medical appointment. This information should be imputed upon the activity scheduling.
- **Persuasion Engine** This engine is used by the two previously presented modules, being its goal create arguments that persuade the users into accept the suggestions of activities and events (Costa et al, 2017a). We believe that the users may be more inclined in accepting certain suggestions if they receive additional information that is related to them. The engine uses three argumentation methods:
 - Argument from analogy;
 - Argument from popular practice;
 - Argument from expert opinion.

The arguments created by these methods are similar to informal speech, to be more relatable to the users. Example: if a *walk in the park* event is suggested, the argument can be *because your doctor has recommended it for lowering your cholesterol*, if the user’s doctor has proposed this activity has health-beneficial. Currently, the arguments are static, but in the next version we aim to implement a natural language engine to be able to create arguments from meta-data and context. In terms of architecture, the persuasion engine is placed between the CLA core modules (the previously explained modules) and the direct messaging system. It has a listener service that receives information from the activities reminder and the events suggestion and accesses its relational databases to find an appropriate argument (according to the received text, to have a proper grammatical sense) that aligns with the internal argumentation methods, and sends it to the direct messaging system. The user response to that argument is saved and if the response is negative other arguments are generated and sent to the user (until a predefined maximum). Furthermore, these responses are saved for future statistical analysis, to improve the engine and recommendation algorithms.

- **Direct Messaging System** This module provides an API (through registered access) that shows the received data directly on the robot touchscreen. This module was specifically designed for the robot, as it uses a screen to relay information, so that other services (external to the CLA) can relay information to the users. The level of access is closed, thus only services that were previously authorized are able to send information. For instance, if the *Emotion Detection* needs to



Fig. 7 Welcoming message.

send a message to the user (if the user need to get close to the camera) it uses this module to send that information. Furthermore, the module is able to capture user's decisions (through touchscreen interaction) using them to establish an interactive experience with the user. Thus, this module is the visual portal to the system and essential to humanize the robot, by keeping an interaction with the users.

In terms of implementation, the CLA is divided in two parts: server and client.

Server The server is implemented with Java micro-webservices that implement the user identification, events suggestion, activities reminder, persuasion engine. The server takes care of the heavy-lifting part, being able to withstand multiple concurrent connections from several users.

Client The client is implemented in Javascript and has the following concurrent services implemented, user identification listener, events listener, activities listener, persuasion listener, direct messaging listener and user response talker. These services run in parallel and have buffering ability, so it is able to show information to the user without losing any data. The way that the client operates is by constant querying the server for new information, in case there is, that information is properly formatted and displayed in the touchscreen. When user interaction is required, the client sends that information to the server and stays waiting for a response. The client's services are non-blocking and multi-threaded.

In Figures 7, 8, 9, 10 we are able to observe the different states of the visual interface. Fig. 7 shows the initial state of the robot, when none is detected or is at a resting state. Fig. 8 shows a kind salute to the user, notifying the user that him/her has detected. Fig. 9 shows an event to the user. This event is presented when the user is interacting with the robot or if the user calendar has no activities at the time. This feature can be configured to suggest one event per day or multiple per day. Fig. 10 shows the Persuasion Engine at work. If the user refuses the suggestion of event/activity, multiple arguments (amount configurable) can be presented to the user to positively nudge the user into accepting.

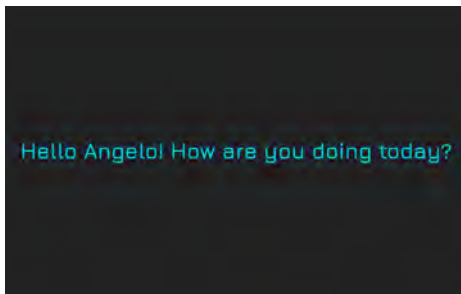


Fig. 8 User identified and saluted.

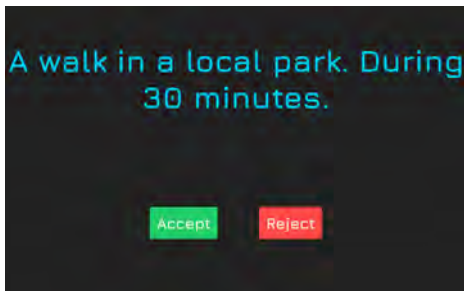


Fig. 9 Event suggested to the user.

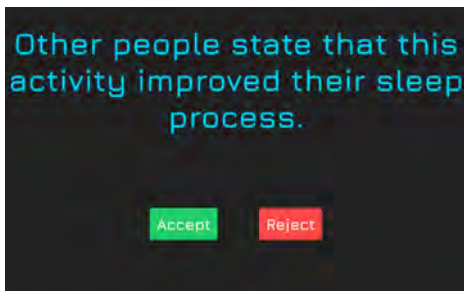


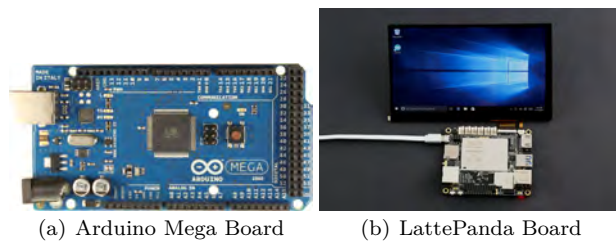
Fig. 10 Argument presented to the user after refusing event in Fig. 9.

4 Hardware Description

This section describes in detail the different types of hardware used to build the robot. The robot was divided into two levels (low level control and high level control) which are described next. The low level control was built using a mega Arduino mega 2560³. This Arduino has a task that controls the different engines and acquires the measurement through sensors. The high level control was built using a Latte-Panda⁴ board. This development card has the operating system that in this case is Windows 10 (Figure 11).

³ <http://www.arduino.cc/>

⁴ <http://www.lattepanda.com/>

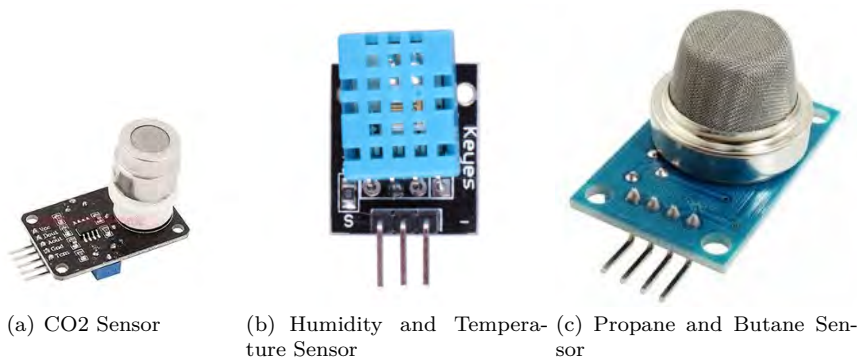


(a) Arduino Mega Board

(b) LattePanda Board

Fig. 11 Developer Boards.

To control the robot's movements, the robot has six servomotors that control the movement of the neck and mouth and a two-step motor that controls the eyes. To access environmental information the robot has a series of sensors (Figure 12) (CO₂, Humidity, Temperature, Propane and Butane) that can help in the installation process, at the same time, the robot uses these sensors at a level of sequences. It means that the robot can use this information to trigger different alarms, such as higher CO₂, temperature, butane gas leakage or propane gas.

(a) CO₂ Sensor

(b) Humidity and Temperature Sensor

(c) Propane and Butane Sensor

Fig. 12 Environmental sensors.

At the same time, these measurements provided by the sensors are sent to the LattePanda board, to be used as an element in the recommendation process. The LattePanda board has connected four elements: camera, microphone, touch screen and speaker which allow the robot to interact with the user. The camera is used to identify the persons and the emotions through image processing as well as a tracking of the person's face. By locating the person's face, the robot can follow his/her face by moving the robot's head. This is very useful when we want to classify the emotion and it is necessary to follow the user's faces to try to detect the emotion. To do this, the coordinates, where the person's face is, are sent to the Arduino board that uses them to move the robot's neck.

The microphone is another way of interacting with the robot. It allows to capture the different voice commands used to wake up the robot or to ask a question and to convert the speech to text. The robot's suggested activities are shown on the touch screen, which the user can use to accept or reject the recommendation. Figure 13

shows the robot, indicating its different parts from the hardware perspective, as well as a frontal view of it.

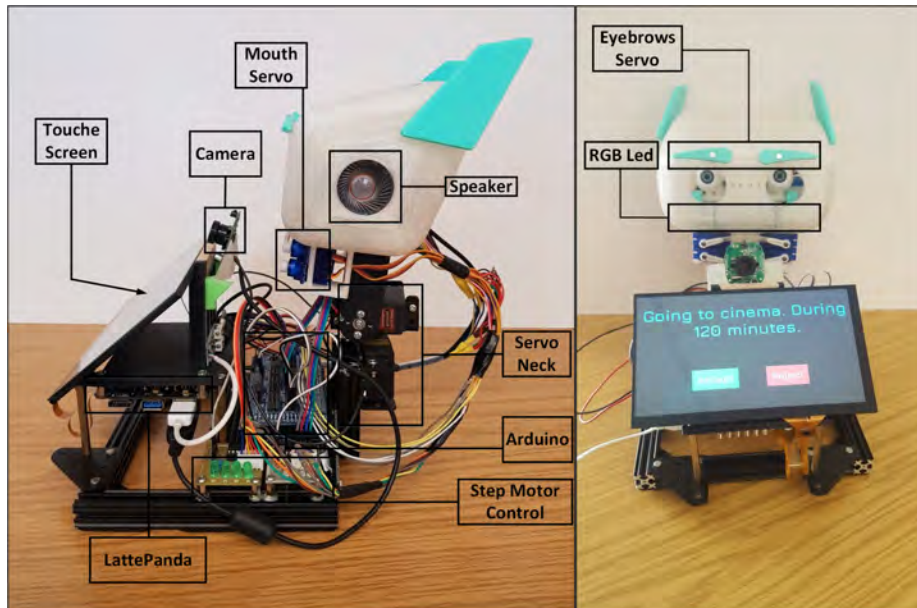


Fig. 13 Robot Parts

5 Robot Working Example

In order to illustrate the behavior of the robot, this section presents an example which describes the different processes made by the robot as an emotional-based persuasive recommender. In the proposed example, the robot will interact with humans by identifying them, detecting their emotions and suggesting activities.

Figure 14 shows the different processes followed by the robot for supporting the described interaction.

In a general view, the robot will follow the following steps:

1. At the beginning, the robot is waiting until a person, in this case can be an elderly person, calls the robot by its name. The name of the root in this example is *EmIR (Emotional Interactive Robot)*. As soon as the robot hears its name, the robot becomes operational.
2. Once awakened, the robot captures a picture through its camera and tries to search for a face in the captured picture. If the robot detects a face, it extracts the coordinates of the face (x, y) . These coordinates are used by the robot in order to track the user. Concretely, the tracking process is carried out through two servomotors located in the neck of the robot. In this way, the camera is centered relatively to the face. Then, the robot takes snapshots periodically, storing it

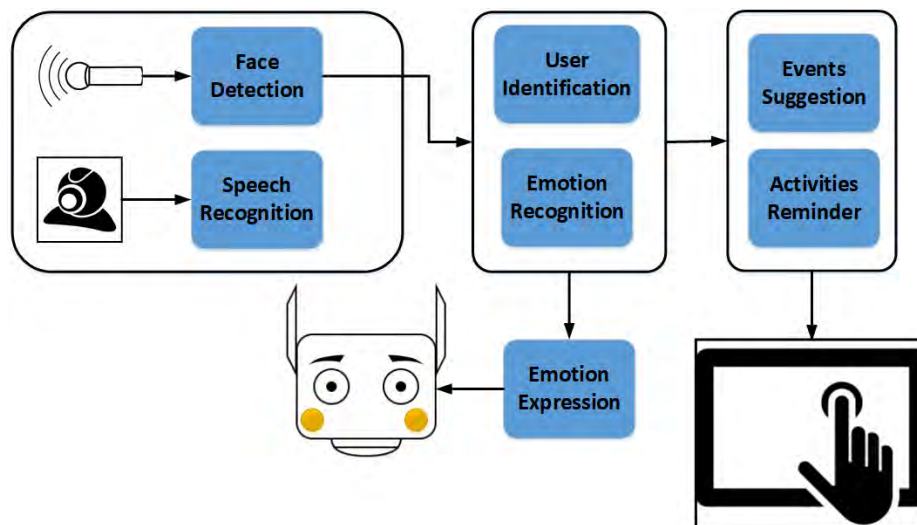


Fig. 14 A schema of the main processes followed by the robot

automatically. Each face image is coded as a vector. The corresponding vector is saved as the identifier of each person and can be enriched with more information such as preferences, gender, age, etc.

3. When the robot registers a new image, it calculates the distance between stored and new vectors. If the distance between these vectors is within the threshold, it means that the person was successfully identified, saluting him/her with his/her name. On the contrary, if the threshold is exceeded, the person is unknown, thus the robot initiates the following dialogue (as an example):
 - **EmIR:** I'm sorry, but I don't know you.
 - **EmIR:** What's your name?.
 - **User:** Hi, my name is Jaime.
4. The robot extracts the person's name which is stored together with the image vector. Once the name is stored, the robot greets the person and detects his/her emotion. The image is then introduced as input into a neural convolutional network that returns the emotional state. If the robot detects more than one face in the same image, the process is repeated for each face, activating the computation of social emotion.
5. Once the emotional state of the user is obtained, the robot will adapt its emotion in an attempt to show empathy. To do this the robot can move its eyebrows and change the lights of its cheeks.
6. All this low-level control of head, eyes, mouth movements and environmental sensors is done through an Arduino. It is in charge of all the environmental measurements done by the different sensors integrated in the robot (such as CO₂, butane and propane sensors). This information can be used by the robot to show alarm messages to the user.
7. Next, the robot makes use of its persuasive recommender system. It can recommend the following items:

- Events that promote active aging. The suggestion of events follows a set of configurations (age, likes, medical condition, etc.) that make them appropriate for each user.
- Activities on the user calendar, e.g., medication reminders, medical appointments.

During this recommendation process, a dialogue with the user is started.

8. If the user refuses an event/activity a set of arguments is generated to persuade the user to accept the recommendation. This process has been explained in detail in section 3.3.
9. After user acceptance or reached the maximum threshold of refuses the robot presents a message saying "Thank you" or other joyful messages. The robot goes then to a waiting state.
10. The robot stays in a waiting state until the persuasive recommender system becomes active again (time configurable). In the meanwhile the robot presents practical information like the current weather state or news, in an effort to interact with the user(s).

6 Conclusions and future work

This paper has presented the EmIR robot, which main goal is to provide assistance to an elderly community. The robot recommends activities to users and also makes reminders about activities already scheduled. Moreover, the robot has been improved presenting human-like features like perceiving emotions of a group of people and displaying human identifiable emotions. The functionalities of the proposed robot are divided into three main aspects. The first one is the Emotion Detection that allows the robot to estimate the emotional state of the people that is in from of the robot. The second one is the Emotion Display, that allows the robot to express empathy with people according to the emotional states previously detected. The third and last one is the User Interaction, that allows the robot to enhance the user experience suggesting activities/events to the user based on his/her profile and medical condition and trying to persuade the users into accepting the suggested activities/events.

The scientific motivation of this work is focused on trying to solve the problem of accompaniment associated with the elderly. This new technology needs to be introduced because of the growing population of older people living alone. These new tools create a new research problem, such as detecting emotions, customizing activities and creating better referrers.

On the market you can find different types of robots used to accompany elderly people. One of the problems of these commercial robots is their high cost. This is why trying to make a qualitative comparison is difficult. This work took as a base some of the characteristics of these commercial robots and added some that did not have like the detection of emotions, environmental sensors and personalized recommendation of activities.

This work is being validated by workers and patients of a daycare centre in the northern area of Portugal. Specifically in the centre *Centro Social Irmandade de S. Torcato*. The validation is being performed through simple interactions with the patients under the supervision of caregivers. Preliminary results show that the robot is well accepted by the elderly people resident of the center.

As future work, we want to improve the detection of emotions by introducing speech recognition as another input in the process of identifying emotional states. Another aspect to be introduced, and currently supported by CAL, is the direct connection with caregivers or doctors in order to share information on scheduled activities or events as well as to share the patient progress. Finally, another important element to take into account is the adaptation of the proposed system to people with disabilities who cannot make use of the touch screen.

Acknowledgements This work is partially supported by the MINECO/FEDER TIN2015-65515-C4-1-R and the FPI grant AP2013-01276 awarded to Jaime-Andres Rincon. This work is supported by COMPETE: POCI-01-0145-FEDER-007043 and FCT - Fundacao para a Ciencia e Tecnologia within the projects UID/CEC/00319/2013 and Post-Doc scholarship SFRH/BPD/102696/2014 (Angelo Costa).

References

- Broekens J, Heerink M, Rosendal H (2009) Assistive social robots in elderly care: a review. *Grontechnology* 8:94–103
- Calvo MG, Lundqvist D (2008) Facial expressions of emotion (kdef): Identification under different display-duration conditions. *Behavior Research Methods* 40(1):109–115, DOI 10.3758/BRM.40.1.109, URL <https://doi.org/10.3758/BRM.40.1.109>
- Caramazza P, Boccolini A, Buschek D, Hullin M, Higham CF, Henderson R, Murray-Smith R, Faccio D (2017) Neural network identification of people hidden from view with a single-pixel, single-photon detector. *CoRR* abs/1709.07244, URL <http://arxiv.org/abs/1709.07244>
- Castro FM, Marín-Jiménez MJ, Guil N, Pérez de la Blanca N (2017) Automatic learning of gait signatures for people identification. In: Rojas I, Joya G, Catala A (eds) *Advances in Computational Intelligence*, Springer International Publishing, Cham, pp 257–270
- Chidambaram V, Chiang YH, Mutlu B (2012) Designing persuasive robots: How robots might persuade people using vocal and nonverbal cues. In: *Proceedings of the Seventh Annual ACM/IEEE International Conference on Human-Robot Interaction*, ACM, New York, NY, USA, HRI '12, pp 293–300
- Costa A, Castillo JC, Novais P, Fernández-Caballero A, Simões R (2012) Sensor-driven agenda for intelligent home care of the elderly. *Expert Syst Appl* 39(15):12,192–12,204, DOI 10.1016/j.eswa.2012.04.058, URL <https://doi.org/10.1016/j.eswa.2012.04.058>
- Costa A, Heras S, Palanca J, Jordán J, Novais P, Julián V (2016) Using argumentation schemes for a persuasive cognitive assistant system. In: *Multi-Agent Systems and Agreement Technologies - 14th European Conference, EUMAS 2016, and 4th International Conference, AT 2016, Valencia, Spain, December 15-16, 2016, Revised Selected Papers*, pp 538–546, DOI 10.1007/978-3-319-59294-7{_}43
- Costa A, Heras S, Palanca J, Jordán J, Novais P, Julián V (2017a) Argumentation schemes for events suggestion in an e-health platform. In: *Persuasive Technology: Development and Implementation of Personalized Technologies to Change Attitudes and Behaviors*, Springer International Publishing, pp 17–30, DOI 10.1007/978-3-319-55134-0{_}2

- Costa A, Novais P, Julian V (2017b) A survey of cognitive assistants. In: Intelligent Systems Reference Library, Springer International Publishing, pp 3–16, DOI 10.1007/978-3-319-62530-0{_}1
- Dynamics I (2018) AIDO. URL <https://www.startengine.com/aido>
- Feil-Seifer D, Mataric M (2005) Socially assistive robotics. in: 9th international conference on rehabilitation robotics. In: 9th International Conference on Rehabilitation Robotics (ICORR 2005), Institute of Electrical and Electronics Engineers (IEEE)
- Ham J, Midden CJH (2014) A persuasive robot to stimulate energy conservation: The influence of positive and negative social feedback and task similarity on energy-consumption behavior. *International Journal of Social Robotics* 6(2):163–171
- Krämer NC (2008) Social effects of virtual assistants. a review of empirical results with regard to communication. In: Prendinger H, Lester J, Ishizuka M (eds) *Intelligent Virtual Agents*, Springer Berlin Heidelberg, Berlin, Heidelberg, pp 507–508
- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*, pp 1097–1105
- Lupiani E, Juarez JM, Palma J, Marin R (2017) Monitoring elderly people at home with temporal case-based reasoning. *Knowledge-Based Systems* 134:116 – 134, DOI <https://doi.org/10.1016/j.knosys.2017.07.025>, URL <http://www.sciencedirect.com/science/article/pii/S0950705117303477>
- Martinez-Martin E, del Pobil AP (2017a) Conflict resolution in robotics: An overview. In: *Advances in Linguistics and Communication Studies*, IGI Global, pp 263–278, DOI 10.4018/978-1-5225-0245-6.ch015
- Martinez-Martin E, del Pobil AP (2017b) Personal robot assistants for elderly care: An overview. In: *Intelligent Systems Reference Library*, Springer International Publishing, pp 77–91, DOI 10.1007/978-3-319-62530-0{_}5
- McCrae RR, John OP (1992) An introduction to the five-factor model and its applications. *Journal of personality* 60(2):175–215, URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1467-6494.1992.tb00970.x/abstract>
- Mehrabian A (1997) Comparison of the PAD and PANAS as models for describing emotions and for differentiating anxiety from depression. *Journal of psychopathology and behavioral assessment* 19(4):331–357, URL <http://link.springer.com/article/10.1007/BF02229025>
- Niemela M, Ylikauppila M, Talja H (2016) Long-term use of paro the therapy robot seal the caregiver perspective. In: *10th World Conference Of Gerontechnology 2016 (ISG 2016)*
- Ortony A (1990) *The Cognitive Structure of Emotions*. Cambridge University Press
- Rincon J, de la Prieta F, Zanardini D, Julian V, Carrascosa C (2017) Influencing over people with a social emotional model. *Neurocomputing* 231:47 – 54, DOI <https://doi.org/10.1016/j.neucom.2016.03.107>, URL <http://www.sciencedirect.com/science/article/pii/S0925231216311602>, *neural Systems in Distributed Computing and Artificial Intelligence*
- Rincon JA, Bajo J, Fernandez A, Julian V, Carrascosa C (2016a) Using emotions for the development of human-agent societies. *Frontiers of Information Technology & Electronic Engineering* 17(4):325–337, DOI 10.1631/FITEE.1500343
- Rincon JA, Costa A, Novais P, Julián V, Carrascosa C (2016b) Using non-invasive wearables for detecting emotions with intelligent agents. In: *International Joint Conference SOCO'16-CISIS'16-ICEUTE'16 - San Sebastián, Spain, October 19th-*

- 21st, 2016, Proceedings, pp 73–84, DOI 10.1007/978-3-319-47364-2{_}8, URL https://doi.org/10.1007/978-3-319-47364-2{_}8
- Siegel M, Breazeal C, Norton MI (2009) Persuasive robotics: The influence of robot gender on human behavior. In: 2009 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp 2563–2568, DOI 10.1109/IROS.2009.5354116
- SoftBank (2018) Pepper. URL <https://www.ald.softbankrobotics.com/en/robots/pepper>
- United Nations, Department of Economic and Social Affairs, Population Division (2015) World population ageing 2015. Tech. Rep. ST/ESA/SER.A/390, United Nations
- United Nations, Department of Economic and Social Affairs, Population Division (2017) World population prospects: The 2017 revision, key findings and advance tables. Tech. Rep. ESA/P/WP/248, United Nations
- Wada K, Shibata T (2008) Social and physiological influences of living with seal robots in an elderly care house for two months. In: The 6th International Conference of the International Society for Gerontechnology
- Wada K, Shibata T, Saito T, Tanie K (2003) Effects of robot assisted activity to elderly people who stay at a health service facility for the aged. In: 3 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp 2847–2852

Biographies



Jaime Andre Rincon was born in Buga (Valle), Colombia, in 1978. He received the B.E. degree in biomedical engineering from the University Manuela Beltran, Colombia, in 2008, and the MS degree in Artificial Intelligence from Universitat Politècnica de València. In 2017 he has received his PhD title in Computer Sciences at the Universitat Politècnica de València. As a researcher his interest is in multi-agent systems, robotics and emotional agents.



Angelo Costas has received his European Doctorate title at the University of Minho (Portugal) in 2013. He has also received his Master's title in 2009 by the University of Minho. Previously he worked as an Invited Assistant Professor at the Polytechnic of Porto, more specifically at the School of Management and Technology. As a researcher his interest is in multi-agent systems, ambient intelligence, mobile computing, machine learning, virtual organizations and privacy and data protection. He has participated in different research projects focused on Ambient Intelligence and Information Analysis both at a European and national level such as AAL4ALL, TIARAC. Additionally, he has been member of the organizing and scientific committee of several international symposiums such as PAAMS, ISAmI, PRO-VE, IDC, etc.



Paulo Novais is an Associate Professor with Habilitation of Computer Science at the Department of Informatics, in the School of Engineering of the University of Minho (Portugal) and a researcher at the ALGORITMI Centre in which he is the coordinator of the research group Synthetic Intelligence Lab, and the coordinator of the research line in "Ambient intelligence for well-being and Health Applications". From the same university he received a PhD in Computer Science in 2003 and his Habilitation in Computer Science in 2011. He was the director of the Master in Informatics Engineering and the Master in Informatics, during 2012/14 and he is the Deputy Director and co-founder of the Master in Law and Informatics at the University of Minho. He has led and participated in several research projects sponsored by Portuguese and European public and private Institutions and has supervised several PhD and MSc students. He is the coauthor of over 170 book chapters, journal papers, conference and workshop papers and books. He was the Vice-president of APPIA (the Portuguese Association for Artificial Intelligence) during 2006/13.



Vicente Julian holds a position of Associate Professor of Computer Science at the Universitat Politècnica de València (UPV) where he has taught since 1996. Vicente Julian is member of the GTI-IA research group, and Deputy Director of the Official Master in Artificial Intelligence, Pattern Recognition and Digital Imaging at the UPV. Four international projects, two international excellence networks, twenty one Spanish projects and four technology transfer projects have covered the research on Artificial Intelligence. He has more than 50 works published in journals with outstanding positions in the list of the Journal Citation Reports, or published in conference proceedings that have a system of external peer review and dissemination of knowledge comparable to journals indexed in relevant positions. Moreover, he has more than 130 contributions and a h-index of 22. Vicente Julian has supervised 8 PhD Thesis.



Carlos Carrascos was born in Valencia (Spain) and received the M.S. degree in Computer Science from the Universidad Politécnica de Valencia (UPV) in 1995. He obtained his Ph.D. in the Departamento de Sistemas Informáticos y Computación at UPV and is currently a lecturer involved in teaching several AI-related subjects at the UPV. His research interests include MAS, social emotions, consensus in MAS, Intelligent Virtual Environments, learning, serious games, information retrieval, and real-time systems.